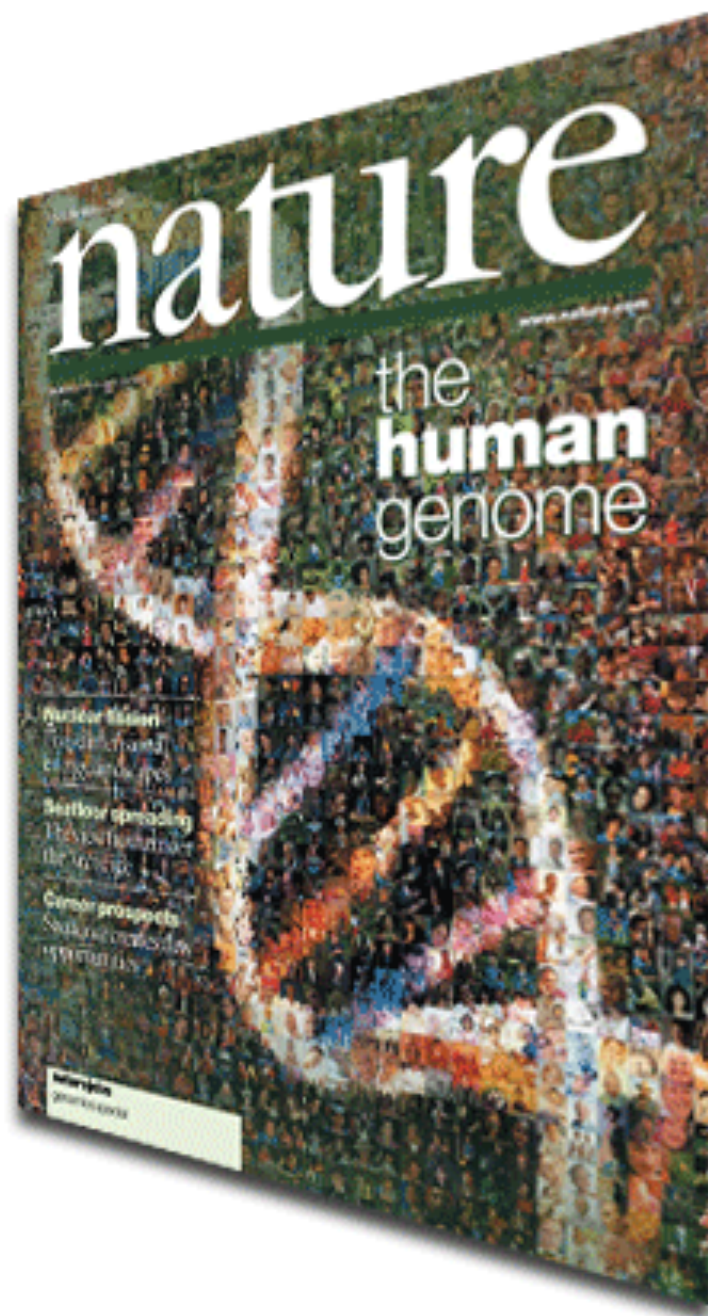
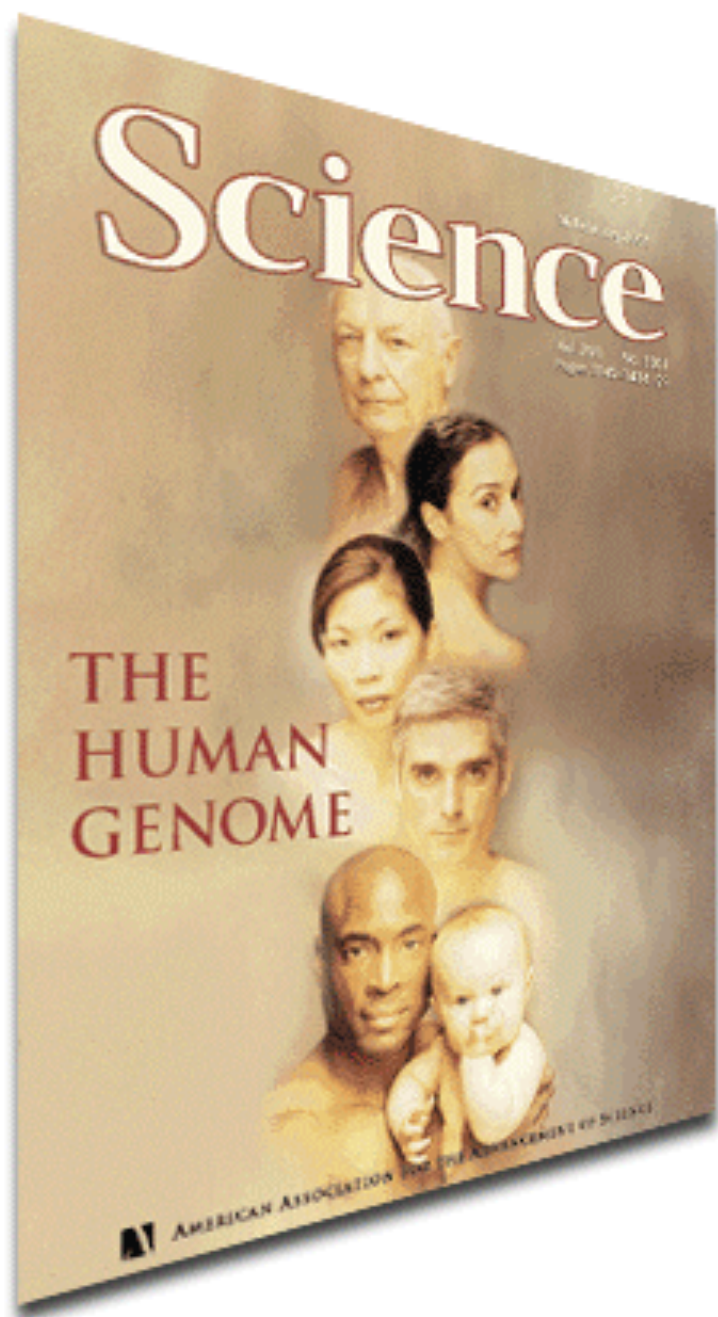


A collage of various mathematical symbols and sequences, including DNA base pairs, Greek letters, and mathematical operators.



Variations: Mutations and Polymorphisms

Single Nucleotide Variation (SNV):

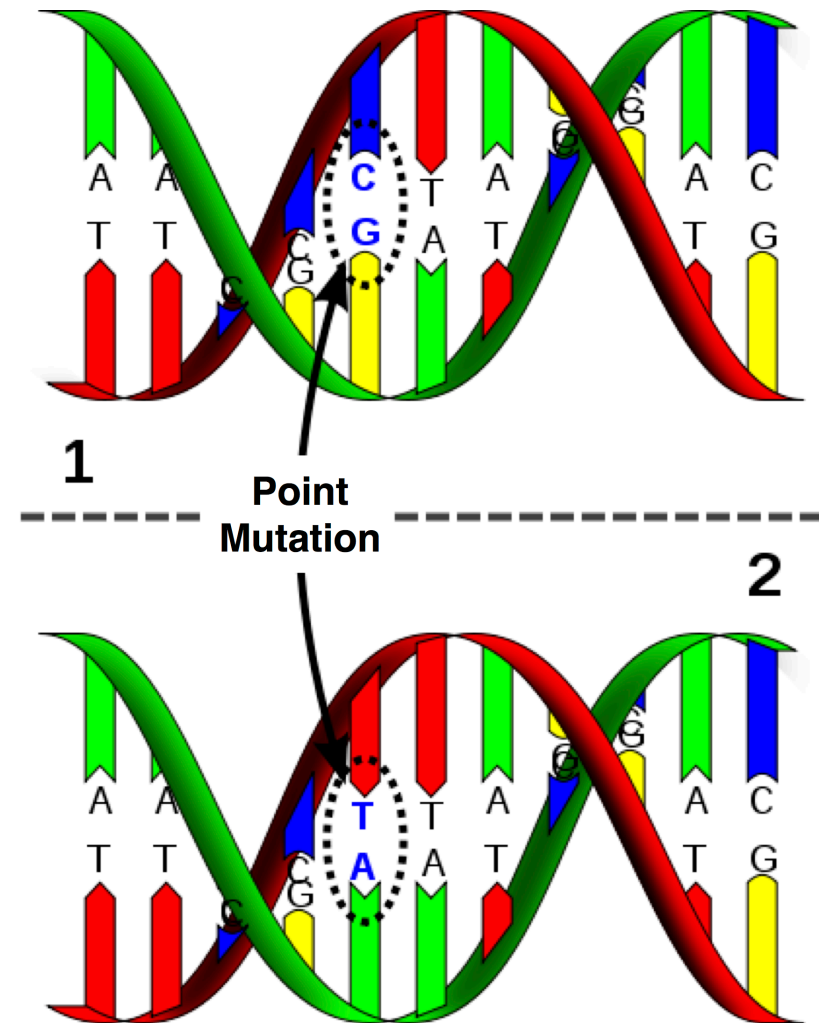
- Permanent change in DNA
- UV, chemicals, tobacco...
- 99.9% identical, 3×10^9 bp

Polymorphism (SNP):

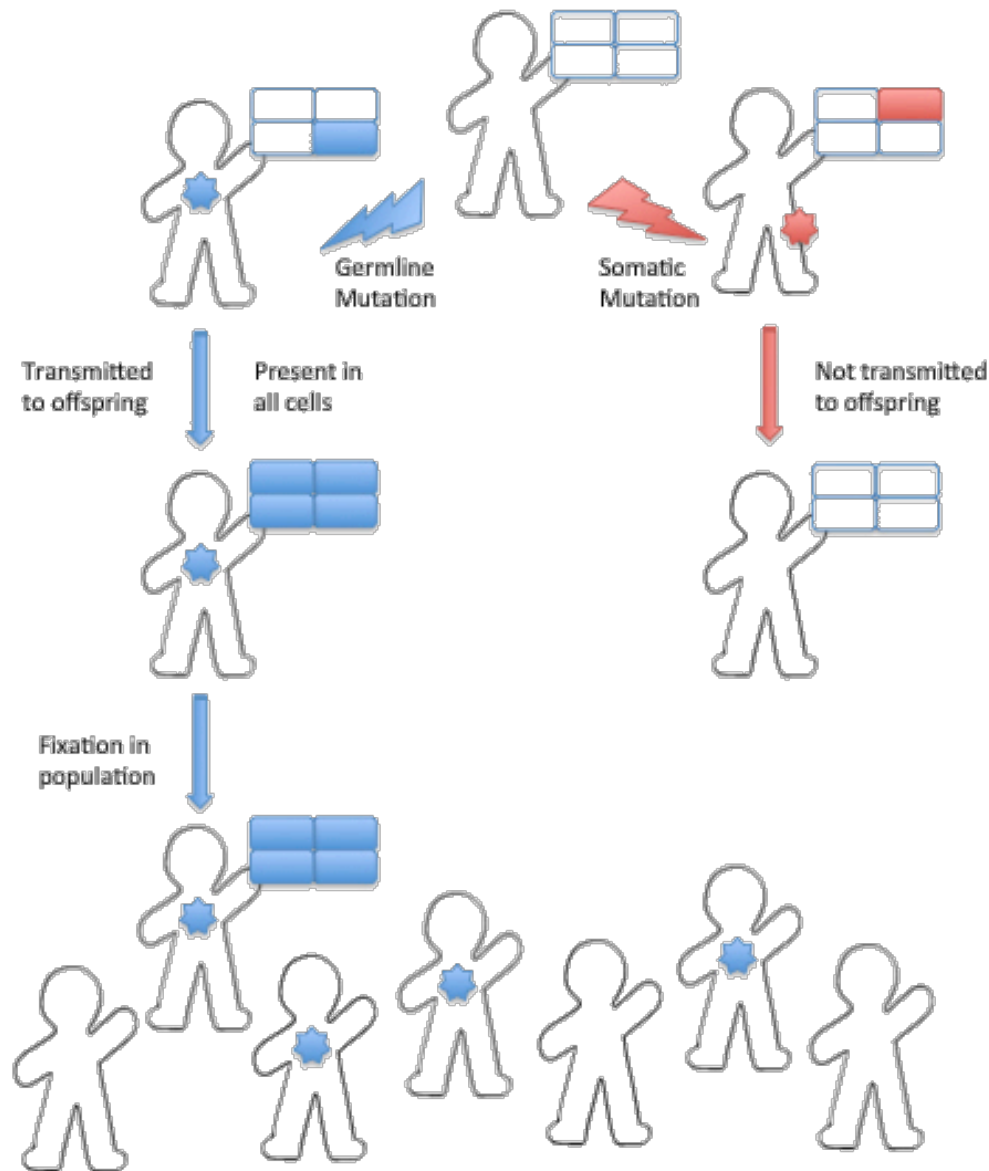
- Implies fixation in population (1%)
- Allele frequencies: Major / Minor

Mutation:

- Usually implies association to disease
- Common allele / Disease allele



Germline vs Somatic Mutations



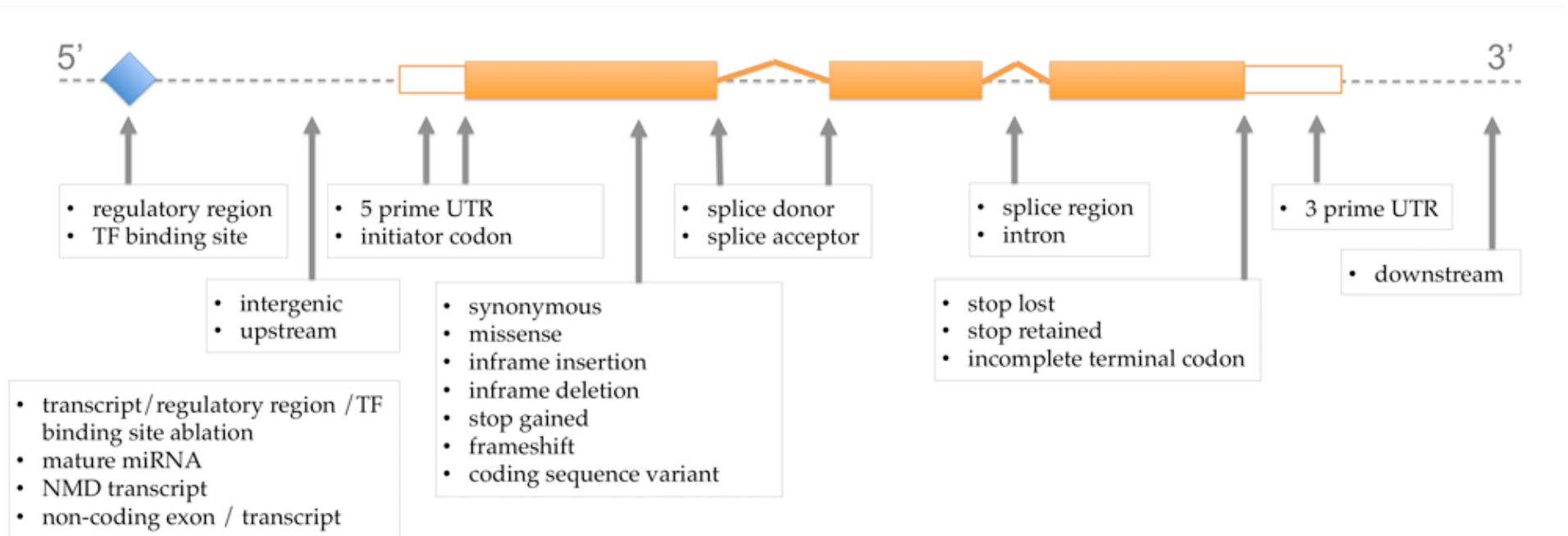
Germline Mutations

- Present in **all** cells
- **Transmitted** to offspring
- **Fixate** in population (SNP)

Somatic Mutations

- Present only in **some** cells
- Not transmitted** to offspring
- Do **not fixate** in population

Coding vs Non-Coding variations



Different regions → Different consequences

- Non-Coding mutations: Regulation, transcription, splicing...
- Coding mutations: non-synonymous, synonymous, stop, frameshift...

Synonymous vs Non-Synonymous Mutations

Degeneration of the genetic code

1) Synonymous mutation

...CCA... → ...CC...

Proline

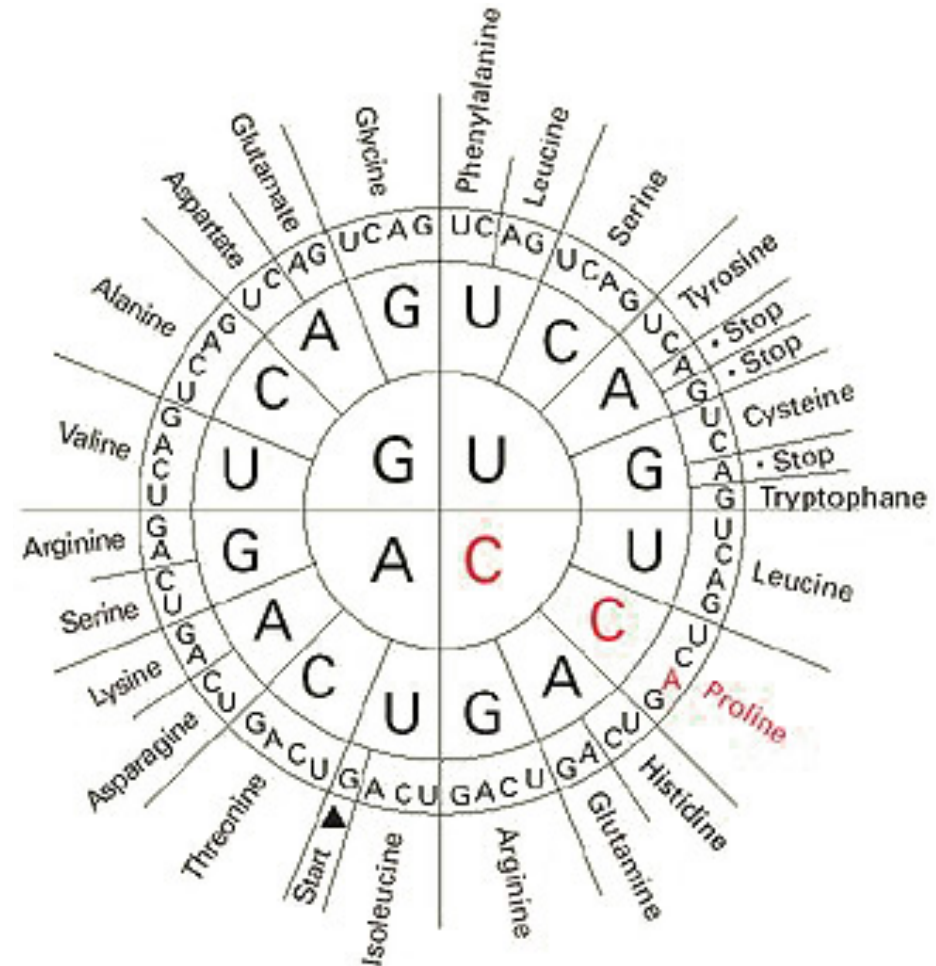
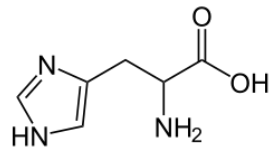
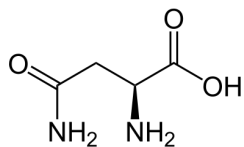
Proline

2) Non-Synonymous mutation

....AAC.... →CAC...

Asparagine

Histidine



Synonymous vs Non-Synonymous Mutations

Degeneration of the genetic code

1) Synonymous mutation

....CCA.... →CCC...

Proline

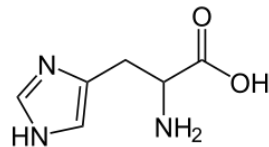
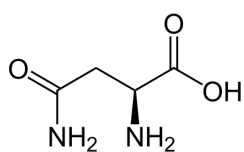
Proline

2) Non-Synonymous mutation

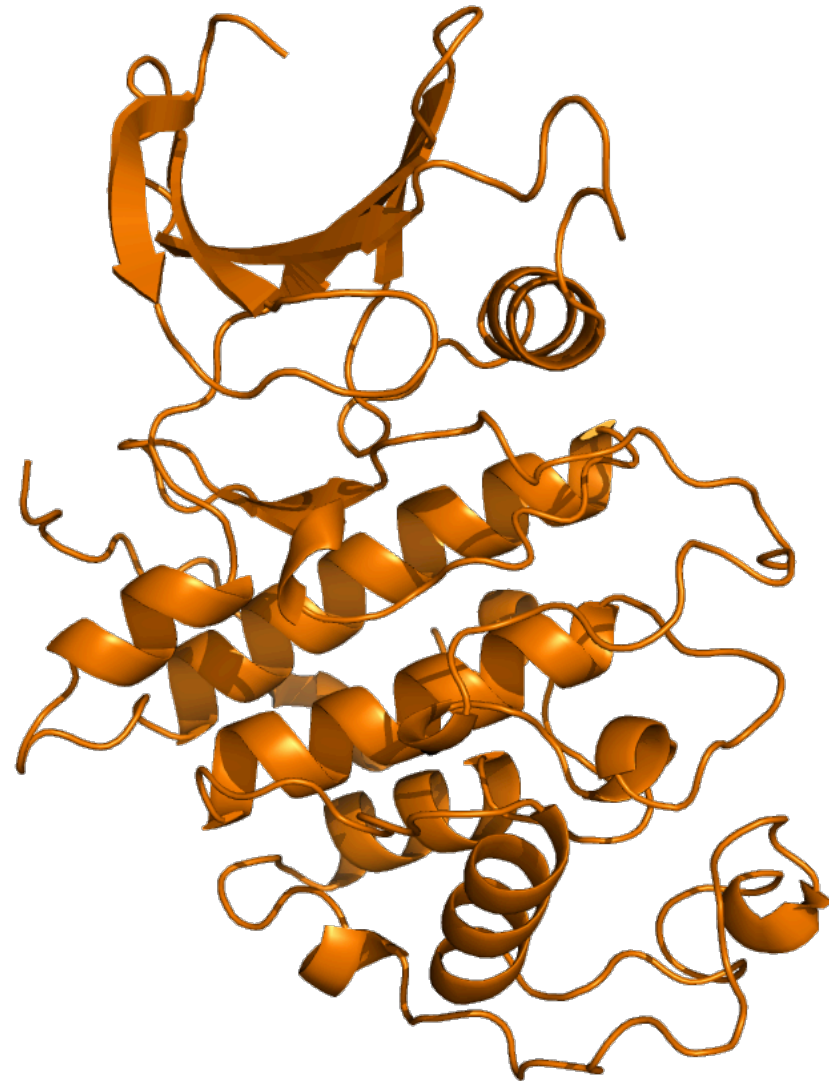
....AAC.... →CAC...

Asparagine

Histidine



- Different chemical properties
- Altered protein structure



Mutations and Disease

[PMID] 18056464 **G 719 S**
 Moreover, known variants, such as epidermal growth factor receptor G719S, that were shown to mediate anticancer drug sensitivity could be detected in other than the previously reported tumor types.
 Induced/natural variant: **0.548738**
 Co-mentioned mutations: G719S
 Breast cancer score: **0.759757**

T790M
G719S

[PMID] 17618705 **T 790 M**
 Recently, it was reported that this acquired resistance is related to a secondary mutation associated with a substitution of threonine by methionine at codon 790 (T790M) of the epidermal growth factor receptor (EGFR) gene.
 Induced/natural variant: **1.07514**
 Co-mentioned mutations: T790M
 Breast cancer score: **2.66059**

[PMID] 17548322 **L 858 R**
 Somatic mutations in TK domain of EGFR were identified in 13 of the 75 (13/75, 17.33%) patients, including 7 cases of in-frame deletion in exon 19 (7/75, 9.33%) and 6 cases of amino acid substitution (2573T>G, L858R) in exon 21 (6/75, 8%) .
 Induced/natural variant: **1.68276**
 Co-mentioned mutations: L858R
 Breast cancer score: **3.19663**

L858R
Y845F

[PMID] 16849545 **Y 845 F**
 Importantly, Cas-dependent protection from the antiproliferative effects of tamoxifen was reversed by the expression of dominant inhibitory variants of these substrates (Y845F EGFR and COOH-terminally truncated STAT5b).
 Induced/natural variant: **-0.30855**
 Co-mentioned mutations: Y845F
 Breast cancer score: **2.20861**

Mutations and Disease

[PMID] 18056464 **G 719 S**
 Moreover, known variants, such as epidermal growth factor receptor G719S, that were shown to mediate anticancer drug sensitivity could be detected in other than the previously reported tumor types.
 Induced/natural variant: **0.548738**
 Co-mentioned mutations: G719S
 Breast cancer score: **0.759757**

Threonine (T) to
Methionine (M) in
position 790

[PMID] 17618705 **T 790 M**
 Recently, it was reported that this acquired resistance is related to a secondary mutation associated with a substitution of threonine by methionine at codon 790 (T790M) of the epidermal growth factor receptor (EGFR) gene.
 Induced/natural variant: **1.07514**
 Co-mentioned mutations: T790M
 Breast cancer score: **2.66059**

[PMID] 17548322 **L 858 R**
 Somatic mutations in TK domain of EGFR were identified in 13 of the 75(13/75, 17.33%) patients, including 7 cases of in-frame deletion in exon 19 (7/75, 9.33%) and 6 cases of amino acid substitution (2573T>G, L858R) in exon 21 (6/75, 8%) .
 Induced/natural variant: **1.68276**
 Co-mentioned mutations: L858R
 Breast cancer score: **3.19663**

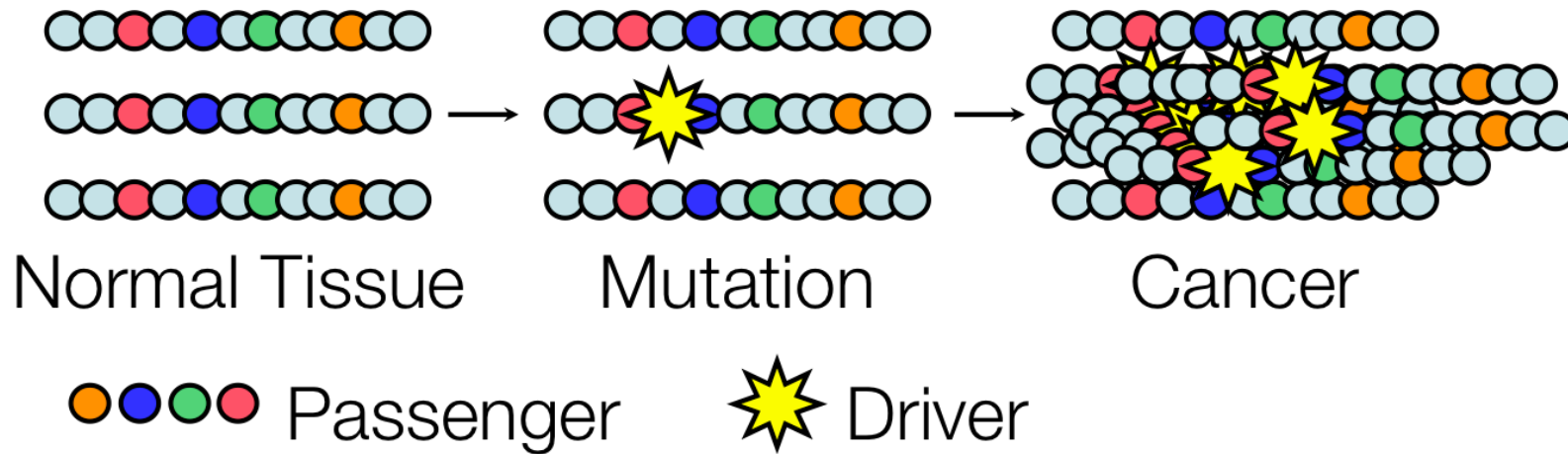
Tyrosine (Y) to
Phenylalanine (F) in
position 845

Y 845 F
 is-dependent protection from the antiproliferative
 xifen was reversed by the expression of dominant
 nts of these substrates (Y845F EGFR and COOH-
 cated STAT5b).
 variant: **-0.30855**
 mutations: Y845F
 Breast cancer score: **2.20861**

The Drivers and Passengers Paradox



The Drivers and Passengers Paradox



- Driver → Confers selective advantage
Disease associated, pathogenic
- Passenger → Present in the clonal progenitor
Functionally neutral



Understanding Mutation in -omics times

Traditionally

1 Mutation
=
1 Disease



Lots of hard work

Phenotype
Function
Mechanism

Now (Bioinformatics)

X Mutations
In
Y Patients
And
Z Conditions



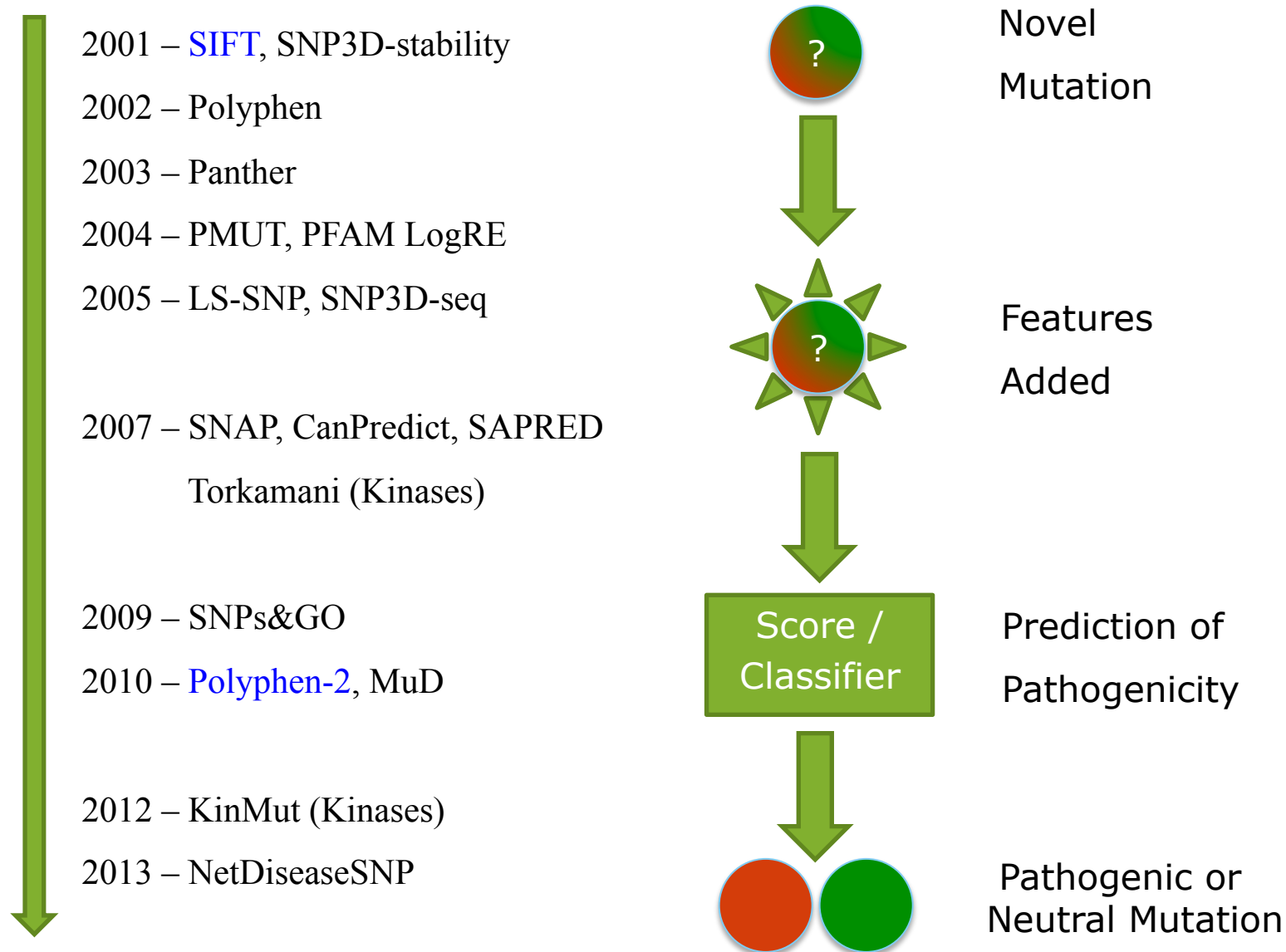
Prediction of
~~Unfeasibility~~ /
Prioritization

SIFT: Pathogenicity of mutations

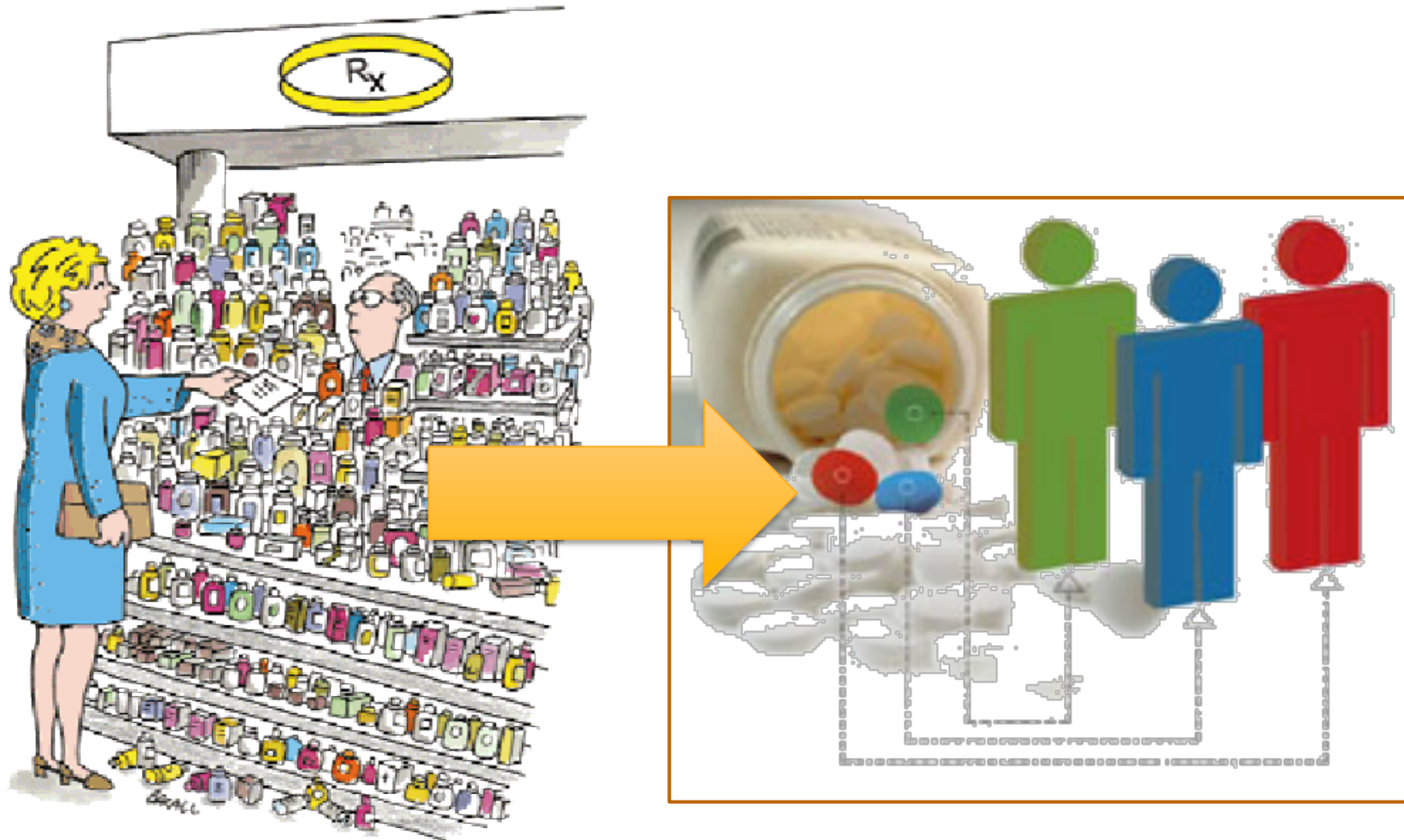
- SIFT is based on amino acid conservation across species
- Mutation of highly conserved → Pathogenic
- <http://blocks.fhcrc.org/sift/SIFT.html>
- Pathogenic if **SIFT score < 0.05**

Human	E	G	R	A	P	D	S	V	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Chimpanzee	E	G	R	A	P	D	S	V	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Orangutan	E	G	R	A	P	D	S	V	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Macaque	E	G	R	A	P	D	S	V	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Mouse	E	G	R	V	P	D	S	I	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Rat	E	G	R	V	P	D	S	I	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Dog	E	S	R	A	P	D	S	V	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Horse	E	G	R	A	P	D	S	I	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Cow	E	G	R	A	P	D	S	V	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Opossum	E	T	K	T	P	D	S	I	D	Y	R	K	K	G	Y	V	T	P	V	K	N
Zebrafish	V	G	K	L	P	K	S	I	D	Y	R	K	L	G	Y	V	T	S	V	K	N
Medaka	L	I	K	L	P	K	S	V	D	Y	R	K	K	G	M	V	T	S	V	K	N
Tetraodon	V	Q	R	L	P	R	N	L	D	Y	R	K	K	G	A	V	T	A	V	K	D
Tetraodon	G	F	E	T	P	P	S	V	D	W	R	K	A	G	L	V	S	P	V	Q	N
Fugu	G	A	D	L	P	Q	T	V	D	W	R	D	K	G	L	V	T	S	V	K	K

Methods to predict pathogenicity



The future: Personalized/Stratified medicine



"Here is my sequence"